

RECEIVED  
CENTRAL FAX CENTER

OCT 18 2006

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE  
BOARD OF PATENT APPEALS AND INTERFERENCES**

In re patent application of:

Kreulen, et al.

Serial No.: 09/629,831

Filed: July 31, 2000

Group Art Unit: 2176

Examiner: Smith, Peter J.

Atty. Docket No.: AM9-99-0157

For: METHOD FOR GENERATION OF AN N-WORK PHRASE DICTIONARY  
FROM A TEXT CORPUS

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Certificate of Transmission by Facsimile

I hereby certify that this correspondence is  
being facsimile transmitted to the United  
States Patent and Trademark Office (Fax No.  
571-273-8300) on 10-18-06

  
Frederick W. Gibb, III

**APPELLANTS' REVISED APPEAL BRIEF**

Sirs:

Appellant respectfully appeals the final rejection of claims 1-17 in the Office Action dated June 14, 2005. A Notice of Appeal was filed on September 16, 2005. A Notice of Non-Compliant Appeal Brief was mailed on September 27, 2006. In response hereto, Appellant is filing a revised "Summary of the Claimed Subject Matter".

Appeal Brief

10/320,318

**I. REAL PARTY IN INTEREST**

The real party in interest is International Business Machines Corp., Armonk, New York, assignee of 100% interest of the above-referenced patent application.

**II. RELATED APPEALS AND INTERFERENCES**

There are no other appeals or interferences known to Appellants, Appellants' legal representative or Assignee which would directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.

**III. STATUS OF CLAIMS**

Claims 1, 6, and 11 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Kostoff et al., hereinafter "Kostoff" (U.S. Patent No. 5,440,481). Claims 2-5, 7-10, and 12-17 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Kostoff and in further view of Kirsch et al., hereinafter "Kirsch" (U.S. Patent No. 6,070,158), Kobayashi (U.S. Patent No. 5,742,834) and Turney (U.S. Patent No. 6,470,307).

**IV. STATUS OF AMENDMENTS**

An After-final Amendment was filed on August 11, 2005. An Advisory Action dated August 23, 2005 indicated that, upon filing an appeal, the Amendment filed on August 11, 2005 did not place the application in condition for allowance, and that the rejections of claims would remain. The claims shown in the appendix are shown in their amended form as of the April 4, 2005 Amendment.

Appeal Brief  
10/320,318

**V. SUMMARY OF CLAIMED SUBJECT MATER**

One feature of the invention is inputting a maximum dictionary size. Claim 1 defines this feature as follows: "inputting a maximum dictionary size." This feature is described at various points in the specification, for example, on page 4 at lines 9 to 15 describes this feature as follows: "The total number of phrases returned will depend upon the user specified maximum dictionary size..., the invention performs a first pass on the set of text documents, as shown in the item 10." This is shown in Figure 1.

Another feature of the invention is determining a frequency of each word in each of the documents. Claim 1 defines this feature as follows: "determining a frequency of each word in each of said documents." This feature is described at various points in the specification, for example page 4, at lines 16 to 18 describes this feature as follows: "Next, in item 11, the invention creates a Hashtable and keeps only the most frequently occurring words in the Hashtable." This is shown in Figure 1.

Another feature of the invention is creating a dictionary of most frequently occurring words in the documents as limited by the maximum dictionary size, such that the dictionary contains less than all words in the documents. Claim 1 defines this feature as follows: "creating a dictionary of most frequently occurring words in said documents as limited by said maximum dictionary size, such that said dictionary contains less than all words in said documents." This feature is described at various points in the specification, for example, on page 4, line 16 to 21 describes this feature as follows: "Next, in item 11, the invention creates a Hashtable and keeps only the most frequently occurring words in the Hashtable. More specifically, the invention finds the V most frequently occurring words in the word-count Hashtable and conserves memory by removing from the Hashtable all words that occur with less frequency than the V most frequently occurring words. Then, as shown in item 12, the invention performs a second pass on the input set of text documents." This is shown in Figure 1.

Another feature of the invention is adding most frequently occurring phrases to the dictionary. Claim 1 defines this feature as follows: "adding most frequently occurring

Appeal Brief

10/320,318

phrases to said dictionary." This feature is described at various points in the specification, for example, on page 4 at lines 24 to Page 5, line 2) describes this feature as follows: "In item 13, the invention adds phrases that are made up only of words in the word-count Hashtable to a phrase-count Hashtable." This is shown in Figure 1.

Another feature of the invention is outputting the most frequently occurring words and the most frequently occurring phrases as the dictionary, such that the dictionary size limits the number of words and phrases maintained in the dictionary. Claim 1 defines this feature as follows: "adding most frequently occurring phrases to said dictionary." This feature is described at various points in the specification, for example, on page 5 at lines 2 to 4 describes this feature as follows: "Finally, in item 14, the invention finds the most frequently occurring V words and phrases in the Hashtables and creates a dictionary of words and phrases from the Hashtables." This is shown in Figure 1.

One feature of the invention is inputting a maximum dictionary size. Claim 6 defines this feature as follows: "inputting a maximum dictionary size." This feature is described at various points in the specification, for example, on page 4 at lines 9 to 15 describes this feature as follows: "The total number of phrases returned will depend upon the user specified maximum dictionary size..., the invention performs a first pass on the set of text documents, as shown in the item 10." This is shown in Figure 1.

Another feature of the invention is determining a frequency of each word in each of the documents. Claim 6 defines this feature as follows: "determining a frequency of each word in each of said documents." This feature is described at various points in the specification, for example page 4, at lines 16 to 18 describes this feature as follows: "Next, in item 11, the invention creates a Hashtable and keeps only the most frequently occurring words in the Hashtable." This is shown in Figure 1.

Another feature of the invention is creating a dictionary of most frequently occurring words in the documents as limited by the maximum dictionary size, such that the dictionary contains less than all words in the documents. Claim 6 defines this feature as follows: "creating a dictionary of most frequently occurring words in said documents as limited by said maximum dictionary size, such that said dictionary contains less than

Appeal Brief

10/320,318

all words in said documents." This feature is described at various points in the specification, for example, on page 4, line 16 to 21 describes this feature as follows: "Next, in item 11, the invention creates a Hashtable and keeps only the most frequently occurring words in the Hashtable. More specifically, the invention finds the V most frequently occurring words in the word-count Hashtable and conserves memory by removing from the Hashtable all words that occur with less frequency than the V most frequently occurring words. Then, as shown in item 12, the invention performs a second pass on the input set of text documents." This is shown in Figure 1.

Another feature of the invention is adding most frequently occurring phrases to the dictionary. Claim 6 defines this feature as follows: "adding most frequently occurring phrases to said dictionary." This feature is described at various points in the specification, for example, on page 4 at lines 24 to Page 5, line 2) describes this feature as follows: "In item 13, the invention adds phrases that are made up only of words in the word-count Hashtable to a phrase-count Hashtable." This is shown in Figure 1.

Another feature of the invention is outputting the most frequently occurring words and the most frequently occurring phrases as the dictionary, such that the dictionary size limits the number of words and phrases maintained in the dictionary. Claim 6 defines this feature as follows: "adding most frequently occurring phrases to said dictionary." This feature is described at various points in the specification, for example, on page 5 at lines 2 to 4 describes this feature as follows: "Finally, in item 14, the invention finds the most frequently occurring V words and phrases in the Hashtables and creates a dictionary of words and phrases from the Hashtables." This is shown in Figure 1.

One feature of the invention is inputting a maximum dictionary size. Claim 11 defines this feature as follows: "inputting a maximum dictionary size." This feature is described at various points in the specification, for example, on page 4 at lines 9 to 15 describes this feature as follows: "The total number of phrases returned will depend upon the user specified maximum dictionary size..., the invention performs a first pass on the set of text documents, as shown in the item 10." This is shown in Figure 1.

## Appeal Brief

10/320,318

Another feature of the invention is determining a frequency of each word in each of the documents. Claim 11 defines this feature as follows: "determining a frequency of each word in each of said documents." This feature is described at various points in the specification, for example page 4, at lines 16 to 18 describes this feature as follows: "Next, in item 11, the invention creates a Hashtable and keeps only the most frequently occurring words in the Hashtable." This is shown in Figure 1.

Another feature of the invention is creating a dictionary of most frequently occurring words in the documents as limited by the maximum dictionary size, such that the dictionary contains less than all words in the documents. Claim 11 defines this feature as follows: "creating a dictionary of most frequently occurring words in said documents as limited by said maximum dictionary size, such that said dictionary contains less than all words in said documents." This feature is described at various points in the specification, on page 4 at lines 16 to 21 describes this feature as follows: "Next, in item 11, the invention creates a Hashtable and keeps only the most frequently occurring words in the Hashtable. More specifically, the invention finds the V most frequently occurring words in the word-count Hashtable and conserves memory by removing from the Hashtable all words that occur with less frequency than the V most frequently occurring words. This is shown in Figure 1.

Another feature of the invention is adding most frequently occurring phrases to the dictionary. Claim 11 defines this feature as follows: "adding most frequently

Best Available Copy